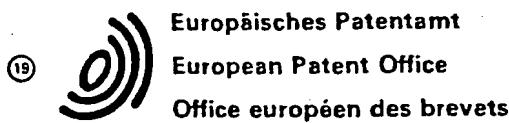


38



(11) Publication number:

0 079 739  
A2

(12)

## EUROPEAN PATENT APPLICATION

(21) Application number: B2305926.6

(51) Int. Cl.3: C 12 N 15/00

C 12 N 1/00, C 12 P 21/02

(22) Date of filing: 08.11.82

C 07 H 21/04, C 07 C 103/52

//C12R1/19, C12R1/865

(30) Priority: 12.11.81 US 320632

(71) Applicant: THE UPJOHN COMPANY  
301 Henrietta Street  
Kalamazoo, Michigan 49001(US)

(43) Date of publication of application:  
25.05.83 Bulletin 83/21

(72) Inventor: Dugaliczyk, Achilles  
c/o The Upjohn Company 301 Henrietta Street  
Kalamazoo Michigan 49001(US)

(84) Designated Contracting States:  
BE CH DE FR GB IT LI NL SE

(74) Representative: Perry, Robert Edward et al,  
GILL JENNINGS & EVERY 53-64 Chancery Lane  
London WC2A 1HN(GB)

(54) Albumin-based nucleotides, their replication and use, and plasmids for use therein.

(55) The DNA sequence coding for human serum albumin has been isolated and inserted as two fragments into two novel plasmids which can be replicated in *E. coli*. These novel fragments can be joined to provide a unitary DNA sequence which then can be cloned into a suitable host, e.g. *E. coli*, for the expression of human serum albumin (which is used extensively in medical practice in treating shock conditions).

EP 0 079 739 A2

ALBUMIN-BASED NUCLEOTIDES, THEIR REPLICATION  
AND USE, AND PLASMIDS FOR USE THEREIN

This invention relates to nucleotides related to human serum albumin (HSA), their replication and use, and plasmids (and host substances) for use therein.

The gene for serum albumin is regulated in 5 development. On the other hand, serum albumin is synthesised in mammals by the adult liver, and its plateau in adulthood. The embryonic liver and yolk sac, on the other hand, produce predominantly  $\alpha$ -fetoprotein, but the synthesis decreases drastically after birth. Recently, 10 Law et al determined the complete sequence of mouse  $\alpha$ -fetoprotein mRNA, Nature 291 (1981) 201-205. The structure revealed extensive homology to mammalian serum albumin, indicating that the two proteins are encoded in the same gene family. Similar conclusions have been 15 reached from studies on the  $\alpha$ -fetoprotein genes of the rat and the mouse; see Jagodzinski et al, Proc. Natl. Acad. Sci. USA, 78 (1981) 3521-3525, and Gorin et al, J. Biol. Chem. 256 (1981) 1954-1959.

The complete nucleotide sequence of human serum 20 mRNA has been determined from recombinant cDNA clones and from a primer-extended cDNA synthesis on the mRNA template. The sequence comprises 2,078 nucleotides, starting upstream of a potential ribosome binding site in the 5'-untranslated region. It contains all the 25 translated codons and extends into the poly(A) at the 3'-terminus. Part of the translated sequence codes for a hydrophobic prepeptide met-lys-trp-val-thr-phe-ile-ser-leu-leu-phe-leu-phe-ser-ser-ala-tyr-ser, followed by a basic propeptide arg-gly-val-phe-arg-arg. These signal 30 peptides are absent from mature serum albumin and, so far, have not been identified in their nascent state in humans. A remaining 1,755 nucleotides of the translated mRNA sequence code for 585 amino acids which are in agreement, with few exceptions, with the published amino 35 acid data for human serum albumin. The mRNA sequence verifies and refines the repeating homology in the triple-domain structure of the serum albumin molecule.

DETAILED DESCRIPTION OF THE INVENTION

Human serum albumin cDNA is cloned into the PstI site of plasmid pBR322 by the oligo(dG)-oligo(dC) tailing technique. Plasmid DNA was isolated from 97 positive colonies which hybridized to the enriched 5 albumin cDNA probe, and the recombinant plasmid pH A36 was found to contain the largest insert of an albumin cDNA sequence. Its restriction endonuclease map is shown in the drawing, together with a restriction map of the primer-extended plasmid clone pH A206. The latter was obtained in a second transformation experiment after initiating 10 the cDNA synthesis from an internal primer. This primer was a 91 base pairs long DNA fragment, MspI(152)-TaqI(182/3), isolated from pH A36. The two plasmids, pH A36 and pH A206, share 0.15 kb of homologous DNA. Together, they encode the entire sequence for human serum albumin, starting with the CTT codon for leu -10 of the prepeptide and extending 15 into the 3'-untranslated region of poly(A).

Sequence of the Albumin cDNA. The sequence was determined for the most part on both DNA strands to ensure accuracy. All of the restriction sites used to end-label DNA fragments were sequenced across by 20 labeling a neighboring restriction site. The entire nucleotide sequence of the serum albumin mRNA, as determined from the cloned DNA in pH A36, pH A206, and from the primer-extended cDNA at the 5'-terminus of the message, is shown in the following Table 1. The inferred amino acid sequence is also indicated. The mRNA length is 2,078 nucleotides, of which 38 represent the 5'-untranslated region, 54 identify a 25 prepeptide of 18 amino acids, 18 identify a propeptide of 6 amino acids, 1,755 code for the known 585 amino acids of serum albumin, 189 make up the 3'-untranslated region and 24 are the poly(A) sequence. Nucleotides 5 to 15 (-34 to -24) in the 5'-untranslated region (Table 30 1) are complementary to a 3'-terminal region of eukaryotic 18S RNA [Azad, A.A. and Deacon, N.J. (1980) Nucl. Acids Res. 8, 4365-4376] and thus could represent a ribosome binding site:

(5')...T T<sup>C</sup> T C T T C T G T.....albumin mRNA  
35 (3')...G A G G A A G G C G U C C m<sub>2</sub><sup>6</sup>A m<sub>2</sub><sup>6</sup>A.....18S RNA

The translated portion of the mRNA sequence codes for the signal peptide and the main body of the albumin polypeptide chain. The

signal peptide is composed of a hydrophobic prepeptide of 18 amino acids and a basic propeptide of 6 amino acids (Table 1). Since pre-peptides are removed from nascent secretory proteins (like albumin) in the endoplasmic reticulum, they are seen only in vitro in heterologous 5 translation systems. As yet, they have not been found within cells [Judah, J.D. and Quinn, P.S. (1977) FEBS 11th Mtg., Copenhagen 50, 21-29; and Strauss, A.W., Donohue, A.M., Bennett, C.D., Rodkey, J.A. and Alberts, A.W. (1977) Proc. Natl. Acad. Sci. USA 74, 1358-1362]. This is the first report of the presence and the sequence of a 10 pre-peptide for human serum albumin. As it is with other secretory proteins, the conversion of proalbumin to albumin takes place in the Golgi vesicles, and the enzyme responsible for this cleavage is probably cathepsin B [Judah, J.D. and Quinn, P.S. (1978) Nature 271, 384-385]. This is also a first report on the sequence of the pro- 15 peptide for normal human serum albumin.

At the 3'-end of the message, the putative polyadenylation signal sequence, AATAAA, is located 164 nucleotides downstream from the amino acid termination codon TAA and 16 nucleotides upstream from the beginning of the poly(A) sequence. Another characteristic sequence 20 located near the polyadenylation site has been identified by Benoist, et al. [Benoist, C., O'Hare, K., Breathnach, R. and Chambon, P. (1980) Nucl. Acids Res. 8, 127-142]; the consensus sequence from several mRNAs was concluded as TTTTCACTGC. A similar sequence, TTTTCTCTGT, is located 19 nucleotides upstream from the AATAAA hexanucleotide in the 25 human albumin mRNA (Table 1).





0079739  
4083

Following are examples which illustrate procedures, including the best mode, for practicing the invention. These examples should not be construed as limiting. All percentages are by weight and all solvent mixture proportions are by volume unless otherwise noted.

5   Example 1      Isolation of Messenger RNA

Human liver mRNA was obtained following the procedure of Chirgwin, et al [Chirgwin, J.M., Przybyla, A.E., MacDonald, R.J. and Rutter, W.J. (1979) Biochemistry 18, 5294-5299]. Immunoprecipitation of albumin containing polysomes was performed according to Taylor and 10 Tse [Taylor, J.M. and Tse, T.P.H. (1976) J. Biol. Chem. 251, 7461-7467]. In vitro translation of mRNA was carried out in a reticulocyte cell-free system, following the instruction of the manufacturer (New England Nuclear). The translation products were separated electrophoretically according to Laemmli [Laemmli, J.K. (1970) Nature 227, 15 680-685.

15   Example 2      Cloning Procedures

Double stranded cDNA was synthesized as described previously [Law, S., Tamaoki, T., Kreuzaler, F. and Dugaiczyk, A. (1980) Gene 10, 53-61]. It was annealed to PstI-linearized pBR322 DNA [Bolivar, F., 20 Rodriguez, R.L., Greene, P.J., Betlach, M.C., Heyneker, H.L., Boyer, H.W., Crossa, J.H. and Falkow, S. (1977) Gene 2, 95-113] that had been tailed with 15 dG residues/3'-terminus [Dugaiczyk, A., Robberson, D.L. and Ullrich, A. (1980) Biochemistry 19, 5869-5873]. The annealed DNA was used to transform E. coli strain RR1, as detailed previously [Law, 25 S., et al., Ibid.]. The albumin clones were selected using the colony hybridization method of Grunstein and Hogness [Grunstein, M. and Hogness, D.S. (1975) Proc. Natl. Acad. Sci. USA 72, 3961-3965], with [<sup>32</sup>P]-labeled cDNA synthesized with the immunoprecipitated polysomal mRNA as template.

30   As shown in Example 5, plasmids pH36 and pH206 were deposited in E. coli HB101 hosts. The plasmids were obtained from E. coli RR1 hosts, described in this example, and transformed into E. coli HR101 by standard procedures well known to those of ordinary skill in this art. The E. coli RR1 hosts were lysed and then centrifuged to 35 separate the chromosomal DNA, cell DNA and plasmid DNA. The plasmid DNA, remaining in the supernatant, is precipitated with ethanol and the precipitate is resuspended in buffer, e.g., TCM (10mM Tris-HCl, pH 8.0, 10 mM CaCl<sub>2</sub>, 10 mM MgCl<sub>2</sub>). The cells for transformation are

prepared as follows: 120 ml of L-broth (1% tryptone, 0.5% yeast extract, 0.5% NaCl) are inoculated with an 18 hour culture of HB101 NRRL B-11371 and grown to an optical density of 0.6 at 600 nm. Cells are washed in cold 100 mM NaCl and resuspended for 15 minutes in 20 ml 5 chilled 50 mM CaCl<sub>2</sub>. Bacteria are then concentrated to one-tenth of this volume in CaCl<sub>2</sub> and mixed 2:1 (v:v) with annealed plasmid DNA, prepared as described above. After chilling the cell-DNA mixture for 15 minutes, it is heat shocked at 42°C for 2 minutes, then allowed to equilibrate at room temperature for ten minutes before addition of 10 L-broth 10 times the volume of the cell-DNA suspension. Transformed cells are incubated in broth at 37°C for one hour before inoculating selective media (L-agar plus 10 µg/ml tetracycline) with 200 µl/plate. Plates are incubated at 37°C for 48 hours to allow the growth of transformants.

15 Example 3 Mapping of Restriction Endonuclease Sites

Restriction endonucleases were obtained from Bethesda Research Laboratories and New England Biolabs and were used according to the manufacturers' instructions. The digested DNA fragments were analyzed electrophoretically on agarose [Helling, R.B., Goodman, H.M. and Boyer, H.W. (1974) J. Virol. 14, 1235-1244] or acrylamide [Dingman, C., Fisher, M.P. and Kakefuda, T. (1972) Biochemistry 11, 1242-1250] gels.

Example 4 DNA Sequencing

DNA fragments were dephosphorylated with bacterial alkaline phosphatase (Worthington) and labeled at the 5'-ends with polynucleotide kinase (Boehringer-Mannheim) and γ[<sup>32</sup>P]ATP. Following digestion with a second restriction endonuclease and electrophoretic separation of the fragments, DNA sequence determination was done according to the procedure of Maxam and Gilbert [Maxam, A. and Gilbert, W. (1980) Methods Enzym. 65, 499-560] and the degradation products were separated electrophoretically on 0.4 mm acrylamide gels as described by Sanger and Coulson [Sanger, F. and Coulson, R. (1978) FEBS Letters 87, 107-110].

Example 5 Recombinant Plasmids pHA36 and pHA206

35 As disclosed in Example 2, albumin clones were selected by hybridizing to the enriched albumin cDNA probe. Plasmid pHA36 contained the largest insert of an albumin cDNA sequence. Both plasmids pHA36 and pHA206 have been deposited in a viable E. coli host in the

permanent collection of the Northern Regional Research Laboratory (NRRL), U.S. Department of Agriculture, Peoria, Illinois, U.S.A. Their accession numbers in this repository are as follows:

HB101(pHA36) - NRRL B-12551

5 HB101(pHA206) - NRRL B-12550

E. coli HB101 is a known and widely available host microbe. Its NRRL accession number is NRRL B-11371.

NRRL B-12550 and NRRL B-12551 are available to the public. ~~upon the grant of a patent. It should be understood that the availability of these deposits does not constitute a license to practice the subject invention in derogation of patent rights granted with the subject instrument by governmental action.~~

15 E. coli RR1 and E. coli HB101 are known and widely available host microbes. Their NRRL accession numbers are NRRL B-12186 and NRRL B-11371, respectively.

pBR322 is a well known and widely available plasmid. It can be obtained from the following host deposit by standard procedures:

NRRL B-12014 - E. coli RR1 (pBR322).

20 YEpl6 is a well known and widely available yeast episomal plasmid. It can be obtained from the following host deposit by standard procedures:

E. coli HB101 (YEpl6) - NRRL B-12093.

Example 6      Assembly of the Serum Albumin Gene

Assembling the pieces together is a straightforward task of restriction enzymology. There is only one MspI site in the overlapping DNA sequence of the two cDNA clones. Two enzymatic steps of (i) MspI digestion of the two DNAs, followed by (ii) the use of ligase, an enzyme that seals DNA fragments, will give the desired product. Although two other undesired DNA species will also be obtained in the course of this recombination reaction, both of them will differ substantially in size. Thus, separation and isolation of the desired DNA species will be achieved.

The assembled DNA clone can be used to transform two types of cells:

35 (a) Escherichia coli

(b) Saccharomyces cerevisiae

(a) The vector of choice is plasmid pBR322, the same that has

been successfully used for cloning of the two fragmented pieces of the serum albumin cDNA.

(b) In order to transform yeast with the serum albumin structural gene sequence, the DNA must be inserted into one of the existing yeast plasmid vectors. This can be accomplished by taking advantage of the fact that several restriction endonuclease recognition sequences are absent from the cloned serum albumin DNA. Synthetic EcoR1 DNA linkers can be ligated to the DNA fragment containing the serum albumin sequence followed by insertion (ligation) into one of the yeast plasmid vectors, e.g., YEp6, at the EcoR1 cloning site. The fused chimeric plasmid can be used to transform yeast according to an established procedure [Hinnen, A., Hicks, J.B. and Fink, G.R. (1978) Proc. Natl. Acad. Sci. USA, 75, 1929]. YEp6 can be obtained from the NRRL repository, as disclosed *supra*.

15 Example 7 Expression of the Serum Albumin Gene

The main body of the structural gene will be transcribed by the E. coli or yeast enzymes. If little or no albumin is produced with the selected host, then an Escherichia coli promoter DNA sequence carrying an initiation codon, i.e., ATG, can be ligated at the beginning of the serum albumin structural gene. Such elements are known and available, e.g., lac promoter used for the expression of human interferon gene in E. coli [Proc. Natl. Acad. Sci. 77, 5230 (1980)]; source of promoter DNA [Proc. Natl. Acad. Sci. 76, 760 (1979)]. Also, see Nature, Vol. 281, October 18, 1979. It has already been documented that such Escherichia coli promoter sequences function well in the expression of foreign genes in Escherichia coli [Mercereau-Puijalon, O., Royal, A., Cami, B., Garapin, A., Krust, A., Gannon, I. and Kourilsky, P. (1978) Nature 275, 505; and Goeddel, D.V., Kleid, D.G., Bolivar, F., Heyneker, H.L., Yansura, D.G., Grea, R., Hirose, T., Kraszewski, A., Itakura, K., and Riggs, A. (1979) Natl. Acad. Sci. USA 76, 106]. For expression in yeast, see Rose, M., Casadaban, M.J. and Botstein, D. (1981) Proc. Natl. Acad. Sci. USA 78, 2460 and 4466.

Example 8 Screening of Clones Producing Albumin

Immunological methods can be used to detect small amounts of albumin made in a bacterium. Flat disks of flexible polyvinyl are coated with the IgG fraction from an immune serum and the disks are pressed onto an agar plate so that antigen released from an in situ lysed microbial colony can bind to the fixed antibody. The plastic

-11-

disk is then incubated with the same total IgG fraction labeled with radioactive iodine so that other determinants on the bound antigen can in turn bind the iodinated antibody. Radioactive areas on the disk expose X-ray film during autoradiography and thus identify colonies 5 producing the protein which is being screened for. Detailed protocols of this procedure have been published [Broome, S. and Gilbert, W. (1978) Proc. Natl. Acad. Sci. USA, 75, 2746]. The purification of human serum albumin can be accomplished by using procedures well known in the art. For example, procedures disclosed in a chapter by T. 10 Peters: Purification and Properties of Serum Albumin, in: The Plasma Proteins, Putnam, Ed. Academic Press, New York, 1975, can be used.

The work described herein was all done in conformity with physical and biological containment requirements specified in the NIH Guidelines.

15

20

25

30

35

CLAIMS

1. Plasmid pHA36, having a restriction endonuclease pattern as shown in the drawing.

5

2. Plasmid pHA206, having a restriction endonuclease pattern as shown in the drawing.

3. E. coli HB101 (pHA36) having the deposit accession number  
10 NRRL B-12551.

4. E. coli HB101 (pHA206) having the deposit accession number  
NRRL B-12550.

15 5. A microorganism modified to contain a nucleotide sequence coding for the amino acid sequence of human serum albumin; said nucleotide sequence is as follows:

20

25

30

35

0079739

4083

-13-

1 -1 -6 p r o -1  
ser ala tyr ser arg gly val phe arg arg asp ala his lys ser glu val ala his arg phe lys asp leu ala glu asn ala lys  
TCG CCT TAT TCC AGC CGT GTC TTT CGT CCA CAT GCA CAC AAG ACT GAC GTT GCT CAT CGG TTT AAA GAT TCC CAA AAA GAT TTC AAA (170)  
  
21 ala leu val ile ala phe ala gln tyr leu gln gln cys pro phe glu asp his val lys leu val asn glu val thr ala phe ala  
GCC TTG CTG ATT GCT CCT CAG TAT CCT CAG CAG TCA GAT GAC TCA GCT GAA AAT TGT GAC AAA TCA CTT CAT ACC CTT GCA TTA GAC AAA TTA GTC AAA ACT CAA TTT GCA (1260)  
  
51 53 60 62  
lys thr asp val ala asp glu ser ala glu asn cys asp lys ser leu his thr leu phe ala asp lys leu cys thr val ala thr leu  
AAA ACA TGT GTC ATT CCT GAT GAC TCA GCT GCA AAA CAA GAA CCT GGC ACA AAT GAA TCA CTT CAT ACC CTT GTC AAA GAT GAC AAC CCA (450)  
  
81 90 91  
arg glu thr tyr gln glu met ala asp cys cys ala lys gln glu pro gly arg asn glu cys phe leu aln his lys asp asn pro  
CGT GAA ACC TAT CCT GAA ATG GCT GAC TGC TGT GCA AAA CAA GAA CCT GGC ACA AAT GAA TCA TGT AAA TAC TTA TAT (350)  
  
111 120 124  
asn leu pro arg leu val arg pro glu val met cys thr ala phe his asp val met cys thr ala phe his asp lys tyr leu try  
AAC CTC CCC CGA TTG CTG AGA CCA CCT CCT GAT GTC ATT GTC ACT GCT TTT CAT GAC AAT GAA GAG ACA TTT TGC AAA TAC TTA TAT (350)  
  
141 150  
glu ile ala arg his pro tyr phe tyr ala pro glu leu leu phe ala lys arg tyr lys ala ala phe thr glu cys cys aln  
GAA ATT CCT AGA AGA CAT CCT TAC TTT TAT GCC CCC GAA CTC CCT TGC ATT CCT AAA AGC TAT AAA CCT CCT TTT ACA GAA TGT TGC CAA (620)  
  
171 177 180  
ala ala asp lys ala ala cys leu leu pro lys leu asp glu leu arg asp glu ala lys ala ser ser ala lys ala arg leu lys cys  
CCT CCT AAA CCT CCC TCC CTC CTC AAC CTC GAT GAA CCT CGG GAT GAA GGG AAG GCT TCC TCT CCC AAA CAC AAC CTC MAG TGT (710)  
  
201 210 220  
ala ser leu gln lys phe gly glu arg ala phe pro lys ala alu phe ala glu  
GCC AGT CTC CAA AAA TTT GCA GAA AGA GCT TTC AAA CCA TGC GCA GCT CCC CTC AGC CAC AGA TTT GAG TTT GCA GAA (300)

0079739

4083

-14-

231                    240                    245 246                    250                    253  
 val ser lys leu val thr asp leu thr lys val his thr glu cys his gly asp leu glu cys als asp arg als asp leu  
 GTC TCC AAG TTA GTG ACA GAT CTT ACC AAA GTC CAC ACG GAA TGC TGC CAT CGA GAT CTC CTT GAA TGT CCT GAT GAC AGC GCG GAC CTT (890)  
 261                    265                    270                    278 279 280                    289 290  
 ala lys tyr lle cys glu een gln asp ser lle ser ser lys leu lys glu cys qly lys pro leu leu glu lys ser his cys lle  
 CCC AAC TAT ATC TGT GAA AAT CAA GAT TCG ATC TCC ACT AAA CTC CTC AAC GAA TCC TGT CAA AAA CCT CTC TTG GAA AAA TCT GAC AAC TAT GCT (980)  
 291                    300                    310                    316                    320  
 ala glu val glu een asp glu met pro ala asp leu pro ser leu ala asp phe val glu ser lys asp val val oys lys asn tyr ala  
 CCC GAA CTC GAA AAT GAT GAC ATG CCT CCT GAC TTG CCT TCA TTA GCT GAT TTT GAA AGT AAC GAT GTC TGT GAA CCT CTC AGA CTT CCC (1070)  
 321                    330                    340                    350  
 glu ala lys asp val phe leu gly met phe leu tyr ala arg arg his pro asp tyr ser val val leu leu asp leu ala  
 GCA GCA AAG GAT GTC TTC TGC CCC ATG TTT TGT TAT GAA TAT GCA AGC CAT CCT GAT TAC TCT GTC GTC CTG CTG AGA CTT CCC (1160)  
 351                    360 361                    369 370                    380  
 lys thr tyr glu thr thr leu glu lys ala ala asp pro his glu cys tyr ala lys val phe glu phe lys dro leu  
 AAG ACA TAT GAA ACC ACT CTA GAG AAG TGC TGT GCT GCC CCT GAT CCT CAT GAA TGC TAT GCC AAA GTC TTC GAT GAA CCT CCT (1250)  
 391                    390                    392                    400  
 val glu glu pro gln een leu lle lys gln een cys glu leu phe glu qly lys val gln asn ala leu leu val arg  
 GTC GAA GAG CCT CAG AAC ATT AAA ATC AAA AAT TGT GAC CTT TTT GAG CAG CCT GCA GAG TAC AAA TTC CAG AAC CTA GCA AAA GTC GGC AGC AAA TGT TGT AAA CAT (1340)  
 411                    420                    430                    437 438                    440  
 tyr thr lys lys val pro gln val ser arg asn leu qly lys val gln ser lys cys cys lys his  
 TAC ACC AAG AAA GCA CCC CAA GTC TCA ACT CCA ACT CTT GTA GAG GTC TCA AGA AAC CTC AAC CAG TTA TGT TGT AAA CAT (1430)  
 441                    448                    450                    460 461                    470  
 pro glu als lys arg met pro oys als glu asp tyr leu ser val val leu asn gln leu oys val gln lys thr pro val ser  
 CCT GAA GCA AAA AGA ATG CCC TGT GCA GAA GAC TAT CTA TCC GTC GTC AAC CAG TCA GTC CAT GAG AAA AGC CCA GTC AGT (1520)  
 471                    476 477                    480                    490  
 asp arg val thr lys cys thr glu ser leu val asp pro cys phe ser als leu glu val asp glu thr tyr val pro lys  
 GAC AGA GTC ACC AAA TGC TCC ACA GAA TCC TGC CTC AAC ACC CGA CCA TGC TTT TCA GCT GAA GTC GAT GAA ACA TAC GTT CCC AAA (1610)  
 501                    510                    514                    520                    530  
 glu phe asn als glu thr phe thr phe his als asp lle cys thr leu ser glu lys glu arg aln lle lys als leu val  
 GAC TTT AAT CCT GAA ACA TTC ACC TTC CAT GCA GTC GAC AGC AGA CAA ACT GCA CTT TCT GTC (1700)

0079739

4083

-15-

531 glu leu val lys his lys pro lys ala thr lys glu glu gln leu lys ala val met asp asp phe val glu lys cys cys lys  
 GAG CTC GTC AAA CAC ACA AAG CCC ACG GCA GAA CTC AAA GAG CAA CTC AAA GCA ACG GCA GAC TGC TGT GCT GCT TTT GCA GAC AAG TGC TGC (1790)  
 535  
 540 glu leu val lys his lys pro lys ala thr lys glu glu gln leu lys ala ala phe val glu lys cys cys lys  
 GAG CTC GTC AAA CAC ACA AAG CCC ACG GCA GAA CTC AAA GAG CAA CTC AAA GCA ACG GCA GAC TGC TGT GCT GCT TTT GCA GAC AAG TGC TGC (1790)  
 545  
 550  
 555  
 560  
 561 ala asp asp lys glu thr cys phe ala glu glu gln ala ala ser gln ala ala ter  
 GCT GAC GAT AAC GAG ACC TGC TTT GCC GAC GCG GCA ACT CAA GCT GCA ACT CAA GCT GCA GAC TGC TGT GCT GCT TTT GCA GAC AAG TGC TGC (1883)  
 565  
 570  
 575  
 580  
 585  
 590  
 595  
 600  
 605  
 610  
 615  
 620  
 625  
 630  
 635  
 640  
 645  
 650  
 655  
 660  
 665  
 670  
 675  
 680  
 685  
 690  
 695  
 700  
 705  
 710  
 715  
 720  
 725  
 730  
 735  
 740  
 745  
 750  
 755  
 760  
 765  
 770  
 775  
 780  
 785  
 790  
 795  
 800  
 805  
 810  
 815  
 820  
 825  
 830  
 835  
 840  
 845  
 850  
 855  
 860  
 865  
 870  
 875  
 880  
 885  
 890  
 895  
 900  
 905  
 910  
 915  
 920  
 925  
 930  
 935  
 940  
 945  
 950  
 955  
 960  
 965  
 970  
 975  
 980  
 985  
 990  
 995  
 1000  
 1005  
 1010  
 1015  
 1020  
 1025  
 1030  
 1035  
 1040  
 1045  
 1050  
 1055  
 1060  
 1065  
 1070  
 1075  
 1080  
 1085  
 1090  
 1095  
 1100  
 1105  
 1110  
 1115  
 1120  
 1125  
 1130  
 1135  
 1140  
 1145  
 1150  
 1155  
 1160  
 1165  
 1170  
 1175  
 1180  
 1185  
 1190  
 1195  
 1200  
 1205  
 1210  
 1215  
 1220  
 1225  
 1230  
 1235  
 1240  
 1245  
 1250  
 1255  
 1260  
 1265  
 1270  
 1275  
 1280  
 1285  
 1290  
 1295  
 1300  
 1305  
 1310  
 1315  
 1320  
 1325  
 1330  
 1335  
 1340  
 1345  
 1350  
 1355  
 1360  
 1365  
 1370  
 1375  
 1380  
 1385  
 1390  
 1395  
 1400  
 1405  
 1410  
 1415  
 1420  
 1425  
 1430  
 1435  
 1440  
 1445  
 1450  
 1455  
 1460  
 1465  
 1470  
 1475  
 1480  
 1485  
 1490  
 1495  
 1500  
 1505  
 1510  
 1515  
 1520  
 1525  
 1530  
 1535  
 1540  
 1545  
 1550  
 1555  
 1560  
 1565  
 1570  
 1575  
 1580  
 1585  
 1590  
 1595  
 1600  
 1605  
 1610  
 1615  
 1620  
 1625  
 1630  
 1635  
 1640  
 1645  
 1650  
 1655  
 1660  
 1665  
 1670  
 1675  
 1680  
 1685  
 1690  
 1695  
 1700  
 1705  
 1710  
 1715  
 1720  
 1725  
 1730  
 1735  
 1740  
 1745  
 1750  
 1755  
 1760  
 1765  
 1770  
 1775  
 1780  
 1785  
 1790  
 1795  
 1800  
 1805  
 1810  
 1815  
 1820  
 1825  
 1830  
 1835  
 1840  
 1845  
 1850  
 1855  
 1860  
 1865  
 1870  
 1875  
 1880  
 1885  
 1890  
 1895  
 1900  
 1905  
 1910  
 1915  
 1920  
 1925  
 1930  
 1935  
 1940  
 1945  
 1950  
 1955  
 1960  
 1965  
 1970  
 1975  
 1980  
 1985  
 1990  
 1995  
 2000  
 2005  
 2010  
 2015  
 2020  
 2025  
 2030  
 2035  
 2040  
 2045  
 2050  
 2055  
 2060  
 2065  
 2070  
 2075  
 2080  
 2085  
 2090  
 2095  
 2100  
 2105  
 2110  
 2115  
 2120  
 2125  
 2130  
 2135  
 2140  
 2145  
 2150  
 2155  
 2160  
 2165  
 2170  
 2175  
 2180  
 2185  
 2190  
 2195  
 2200  
 2205  
 2210  
 2215  
 2220  
 2225  
 2230  
 2235  
 2240  
 2245  
 2250  
 2255  
 2260  
 2265  
 2270  
 2275  
 2280  
 2285  
 2290  
 2295  
 2300  
 2305  
 2310  
 2315  
 2320  
 2325  
 2330  
 2335  
 2340  
 2345  
 2350  
 2355  
 2360  
 2365  
 2370  
 2375  
 2380  
 2385  
 2390  
 2395  
 2400  
 2405  
 2410  
 2415  
 2420  
 2425  
 2430  
 2435  
 2440  
 2445  
 2450  
 2455  
 2460  
 2465  
 2470  
 2475  
 2480  
 2485  
 2490  
 2495  
 2500  
 2505  
 2510  
 2515  
 2520  
 2525  
 2530  
 2535  
 2540  
 2545  
 2550  
 2555  
 2560  
 2565  
 2570  
 2575  
 2580  
 2585  
 2590  
 2595  
 2600  
 2605  
 2610  
 2615  
 2620  
 2625  
 2630  
 2635  
 2640  
 2645  
 2650  
 2655  
 2660  
 2665  
 2670  
 2675  
 2680  
 2685  
 2690  
 2695  
 2700  
 2705  
 2710  
 2715  
 2720  
 2725  
 2730  
 2735  
 2740  
 2745  
 2750  
 2755  
 2760  
 2765  
 2770  
 2775  
 2780  
 2785  
 2790  
 2795  
 2800  
 2805  
 2810  
 2815  
 2820  
 2825  
 2830  
 2835  
 2840  
 2845  
 2850  
 2855  
 2860  
 2865  
 2870  
 2875  
 2880  
 2885  
 2890  
 2895  
 2900  
 2905  
 2910  
 2915  
 2920  
 2925  
 2930  
 2935  
 2940  
 2945  
 2950  
 2955  
 2960  
 2965  
 2970  
 2975  
 2980  
 2985  
 2990  
 2995  
 3000  
 3005  
 3010  
 3015  
 3020  
 3025  
 3030  
 3035  
 3040  
 3045  
 3050  
 3055  
 3060  
 3065  
 3070  
 3075  
 3080  
 3085  
 3090  
 3095  
 3100  
 3105  
 3110  
 3115  
 3120  
 3125  
 3130  
 3135  
 3140  
 3145  
 3150  
 3155  
 3160  
 3165  
 3170  
 3175  
 3180  
 3185  
 3190  
 3195  
 3200  
 3205  
 3210  
 3215  
 3220  
 3225  
 3230  
 3235  
 3240  
 3245  
 3250  
 3255  
 3260  
 3265  
 3270  
 3275  
 3280  
 3285  
 3290  
 3295  
 3300  
 3305  
 3310  
 3315  
 3320  
 3325  
 3330  
 3335  
 3340  
 3345  
 3350  
 3355  
 3360  
 3365  
 3370  
 3375  
 3380  
 3385  
 3390  
 3395  
 3400  
 3405  
 3410  
 3415  
 3420  
 3425  
 3430  
 3435  
 3440  
 3445  
 3450  
 3455  
 3460  
 3465  
 3470  
 3475  
 3480  
 3485  
 3490  
 3495  
 3500  
 3505  
 3510  
 3515  
 3520  
 3525  
 3530  
 3535  
 3540  
 3545  
 3550  
 3555  
 3560  
 3565  
 3570  
 3575  
 3580  
 3585  
 3590  
 3595  
 3600  
 3605  
 3610  
 3615  
 3620  
 3625  
 3630  
 3635  
 3640  
 3645  
 3650  
 3655  
 3660  
 3665  
 3670  
 3675  
 3680  
 3685  
 3690  
 3695  
 3700  
 3705  
 3710  
 3715  
 3720  
 3725  
 3730  
 3735  
 3740  
 3745  
 3750  
 3755  
 3760  
 3765  
 3770  
 3775  
 3780  
 3785  
 3790  
 3795  
 3800  
 3805  
 3810  
 3815  
 3820  
 3825  
 3830  
 3835  
 3840  
 3845  
 3850  
 3855  
 3860  
 3865  
 3870  
 3875  
 3880  
 3885  
 3890  
 3895  
 3900  
 3905  
 3910  
 3915  
 3920  
 3925  
 3930  
 3935  
 3940  
 3945  
 3950  
 3955  
 3960  
 3965  
 3970  
 3975  
 3980  
 3985  
 3990  
 3995  
 4000  
 4005  
 4010  
 4015  
 4020  
 4025  
 4030  
 4035  
 4040  
 4045  
 4050  
 4055  
 4060  
 4065  
 4070  
 4075  
 4080  
 4085  
 4090  
 4095  
 4100  
 4105  
 4110  
 4115  
 4120  
 4125  
 4130  
 4135  
 4140  
 4145  
 4150  
 4155  
 4160  
 4165  
 4170  
 4175  
 4180  
 4185  
 4190  
 4195  
 4200  
 4205  
 4210  
 4215  
 4220  
 4225  
 4230  
 4235  
 4240  
 4245  
 4250  
 4255  
 4260  
 4265  
 4270  
 4275  
 4280  
 4285  
 4290  
 4295  
 4300  
 4305  
 4310  
 4315  
 4320  
 4325  
 4330  
 4335  
 4340  
 4345  
 4350  
 4355  
 4360  
 4365  
 4370  
 4375  
 4380  
 4385  
 4390  
 4395  
 4400  
 4405  
 4410  
 4415  
 4420  
 4425  
 4430  
 4435  
 4440  
 4445  
 4450  
 4455  
 4460  
 4465  
 4470  
 4475  
 4480  
 4485  
 4490  
 4495  
 4500  
 4505  
 4510  
 4515  
 4520  
 4525  
 4530  
 4535  
 4540  
 4545  
 4550  
 4555  
 4560  
 4565  
 4570  
 4575  
 4580  
 4585  
 4590  
 4595  
 4600  
 4605  
 4610  
 4615  
 4620  
 4625  
 4630  
 4635  
 4640  
 4645  
 4650  
 4655  
 4660  
 4665  
 4670  
 4675  
 4680  
 4685  
 4690  
 4695  
 4700  
 4705  
 4710  
 4715  
 4720  
 4725  
 4730  
 4735  
 4740  
 4745  
 4750  
 4755  
 4760  
 4765  
 4770  
 4775  
 4780  
 4785  
 4790  
 4795  
 4800  
 4805  
 4810  
 4815  
 4820  
 4825  
 4830  
 4835  
 4840  
 4845  
 4850  
 4855  
 4860  
 4865  
 4870  
 4875  
 4880  
 4885  
 4890  
 4895  
 4900  
 4905  
 4910  
 4915  
 4920  
 4925  
 4930  
 4935  
 4940  
 4945  
 4950  
 4955  
 4960  
 4965  
 4970  
 4975  
 4980  
 4985  
 4990  
 4995  
 5000  
 5005  
 5010  
 5015  
 5020  
 5025  
 5030  
 5035  
 5040  
 5045  
 5050  
 5055  
 5060  
 5065  
 5070  
 5075  
 5080  
 5085  
 5090  
 5095  
 5100  
 5105  
 5110  
 5115  
 5120  
 5125  
 5130  
 5135  
 5140  
 5145  
 5150  
 5155  
 5160  
 5165  
 5170  
 5175  
 5180  
 5185  
 5190  
 5195  
 5200  
 5205  
 5210  
 5215  
 5220  
 5225  
 5230  
 5235  
 5240  
 5245  
 5250  
 5255  
 5260  
 5265  
 5270  
 5275  
 5280  
 5285  
 5290  
 5295  
 5300  
 5305  
 5310  
 5315  
 5320  
 5325  
 5330  
 5335  
 5340  
 5345  
 5350  
 5355  
 5360  
 5365  
 5370  
 5375  
 5380  
 5385  
 5390  
 5395  
 5400  
 5405  
 5410  
 5415  
 5420  
 5425  
 5430  
 5435  
 5440  
 5445  
 5450  
 5455  
 5460  
 5465  
 5470  
 5475  
 5480  
 5485  
 5490  
 5495  
 5500  
 5505  
 5510  
 5515  
 5520  
 5525  
 5530  
 5535  
 5540  
 5545  
 5550  
 5555  
 5560  
 5565  
 5570  
 5575  
 5580  
 5585  
 5590  
 5595  
 5600  
 5605  
 5610  
 5615  
 5620  
 5625  
 5630  
 5635  
 5640  
 5645  
 5650  
 5655  
 5660  
 5665  
 5670  
 5675  
 5680  
 5685  
 5690  
 5695  
 5700  
 5705  
 5710  
 5715  
 5720  
 5725  
 5730  
 5735  
 5740  
 5745  
 5750  
 5755  
 5760  
 5765  
 5770  
 5775  
 5780  
 5785  
 5790  
 5795  
 5800  
 5805  
 5810  
 5815  
 5820  
 5825  
 5830  
 5835  
 5840  
 5845  
 5850  
 5855  
 5860  
 5865  
 5870  
 5875  
 5880  
 5885  
 5890  
 5895  
 5900  
 5905  
 5910  
 5915  
 5920  
 5925  
 5930  
 5935  
 5940  
 5945  
 5950  
 5955  
 5960  
 5965  
 5970  
 5975  
 5980  
 5985  
 5990  
 5995  
 6000  
 6005  
 6010  
 6015  
 6020  
 6025  
 6030  
 6035  
 6040  
 6045  
 6050  
 6055  
 6060  
 6065  
 6070  
 6075  
 6080  
 6085  
 6090  
 6095  
 6100  
 6105  
 6110  
 6115  
 6120  
 6125  
 6130  
 6135  
 6140  
 6145  
 6150  
 6155  
 6160  
 6165  
 6170  
 6175  
 6180  
 6185  
 6190  
 6195  
 6200  
 6205  
 6210  
 6215  
 6220  
 6225  
 6230  
 6235  
 6240  
 6245  
 6250  
 6255  
 6260  
 6265  
 6270  
 6275  
 6280  
 6285  
 6290  
 6295  
 6300  
 6305  
 6310  
 6315  
 6320  
 6325  
 6330  
 6335  
 6340  
 6345  
 6350  
 6355  
 6360  
 6365  
 6370  
 6375  
 6380  
 6385  
 6390  
 6395  
 6400  
 6405  
 6410  
 6415  
 6420  
 6425  
 6430  
 6435  
 6440  
 6445  
 6450  
 6455  
 6460  
 6465  
 6470  
 6475  
 6480  
 6485  
 6490  
 6495  
 6500  
 6505  
 6510  
 6515  
 6520  
 6525  
 6530  
 6535  
 6540  
 6545  
 6550  
 6555  
 6560  
 6565  
 6570  
 6575  
 6580  
 6585  
 6590  
 6595  
 6600  
 6605  
 6610  
 6615  
 6620  
 6625  
 6630  
 6635  
 6640  
 6645  
 6650  
 6655  
 6660  
 6665  
 6670  
 6675  
 6680  
 6685  
 6690  
 6695  
 6700  
 6705  
 6710  
 6715  
 6720  
 6725  
 6730  
 6735  
 6740  
 6745  
 6750  
 6755  
 6760  
 6765  
 6770  
 6775  
 6780  
 6785  
 6790  
 6795  
 6800  
 6805  
 6810  
 6815  
 6820  
 6825  
 6830  
 6835  
 6840  
 6845  
 6850  
 6855  
 6860  
 6865  
 6870  
 6875  
 6880  
 6885  
 6890  
 6895  
 6900  
 6905  
 6910  
 6915  
 6920  
 6925  
 6930  
 6935  
 6940  
 6945  
 6950  
 6955  
 6960  
 6965  
 6970  
 6975  
 6980  
 6985  
 6990  
 6995  
 7000  
 7005  
 7010  
 7015  
 7020  
 7025  
 7030  
 7035  
 7040  
 7045  
 7050  
 7055  
 7060  
 7065  
 7070  
 7075  
 7080  
 7085  
 7090  
 7095  
 7100  
 7105  
 7110  
 7115  
 7120  
 7125  
 7130  
 7135  
 7140  
 7145  
 7150  
 7155  
 7160  
 7165  
 7170  
 7175  
 7180  
 7185  
 7190  
 7195  
 7200  
 7205  
 7210  
 7215  
 7220  
 7225  
 7230  
 7235  
 7240  
 7245  
 7250  
 7255  
 7260  
 7265  
 7270  
 7275  
 7280  
 7285  
 7290  
 7295  
 7300  
 7305  
 7310  
 7315  
 7320  
 7325  
 7330  
 7335  
 7340  
 7345  
 7350  
 7355  
 7360  
 7365  
 7370  
 7375  
 7380  
 7385  
 7390  
 7395  
 7400  
 7405  
 7410  
 7415  
 7420  
 7425  
 7430  
 7435  
 7440  
 7445  
 7450  
 7455  
 7460  
 7465  
 7470  
 7475  
 7480  
 7485  
 7490  
 7495  
 7500  
 7505  
 7510  
 7515  
 7520  
 7525  
 7530  
 7535  
 7540  
 7545  
 7550  
 7555  
 7560  
 7565  
 7570  
 7575  
 7580  
 7585  
 7590  
 7595  
 7600  
 7605  
 7610  
 7615  
 7620  
 7625  
 7630  
 7635  
 7640  
 7645  
 7650  
 7655  
 7660  
 7665  
 7670  
 7675  
 7680  
 7685  
 7690  
 7695  
 7700  
 7705  
 7710  
 7715  
 7720  
 7725  
 7730  
 7735  
 7740  
 7745  
 7750  
 7755  
 7760  
 7765  
 7770  
 7775  
 7780  
 7785  
 7790  
 7795  
 7800  
 7805  
 7810  
 7815  
 7820  
 7825  
 7830  
 7835  
 7840  
 7845  
 7850  
 7855  
 7860  
 7865  
 7870  
 7875  
 7880  
 7885  
 7890  
 7895  
 7900  
 7905  
 7910  
 7915  
 7920  
 7925  
 7930  
 7935  
 7940  
 7945  
 7950  
 7955  
 7960  
 7965  
 7970  
 7975  
 7980  
 7985  
 7990  
 7995  
 8000  
 8005  
 8010  
 8015  
 8020  
 8025  
 8030  
 8035  
 8040  
 8045  
 8050  
 8055  
 8060  
 8065  
 8070  
 8075  
 8080  
 8085  
 8090  
 8095  
 8100  
 8105  
 8110  
 811

6. Nucleotide sequence of the cDNA of human serum albumin, said nucleotide sequence is as follows:

-17-

0079739  
4083

0079739

4083

-18-

5

10

15

20

25

30

35

0079739

4083

7. Nucleotide sequence coding for the prepeptide of human serum albumin, said nucleotide sequence is as follows:

5

10

15

20

25

30

35

8. Nucleotide sequence coding for pro human serum albumin, said nucleotide sequence is as follows:

0079739  
4083

-21-

35           38           25           20           15           10           5  
231    ser lys leu val thr asp leu thr lys val his thr glu cys his gly asp leu leu glu cys ala asp arg ala asp leu  
      GTT TCC AAG TTA GTG ACA GAT CTT ACC AAA GTC CAC ACG GAA TGC TGC CAT GCA TGT GCT GAA TGT GAT GAC AGC GCG GAC CTT (890)  
  
261    265           270           278 279 280           289 290  
ala lys tyr ile oys glu asn gln asp ser lle ser aer lys leu lys glu oys cys glu lys pro leu leu glu lys ser his cys lle  
      GCC AAG TAT ATC TGT GAA AAT CAA GAT TCG ATC TCC AGT AAA CTG AAG GAA TCC TGT GAA AAA CCT CTG TTG GAA AAC TCC TGC ATT (980)  
  
291           310           316           320  
ala glu val glu asn asp glu met pro ala asp leu pro ser leu ala ala asp phe val glu ser lys asp val cys lys asn tyr ala  
      GCC GAA GTC GAA AAT GAT GAC ATG CCT GCT GAT CCT GTC TGA TAT GCA TTA GCT GAT TTT GTC GAA AGT AAC GAT GTC TGT AAA AAC TAT GCT (1070)  
  
321           330           340           350  
glu ala lys asp val phe leu gly met phe leu tyr ala asp arg arg his pro asp tyr ser val val phe lys pro leu leu ala  
      GAC GCA AAG GAT GTC TTC TGC TGT GGC ATG TTT TGC TAT GAA TAT GCA ACT CCT GAT CCT GCA GAT CTC TAC TAC GAT TAC TCT GCA CTC AGA CTT GCC (1160)  
  
351           360 361           369 370           380  
lys thr tyr glu thr thr leu glu lys cys cys ala ala asp pro his glu cys tyr ala lys val phe asp glu phe lys pro leu  
      AAC ACA TAT GAA ACC ACT CTA GAG TGC TGT GGC GCT GCA GAT CCT GCA TAA TGC TAT GCA TGT GAA TTT AAA CCT CCT CTC CGT (1250)  
  
381           390           392           400           410  
val glu glu pro gln asn leu lle lys gln asn cys glu leu phe glu qln leu qly glu tyr lys phe gln asn ala leu leu val arg  
      GTC GAA GAG CCT CAG AAC TTA ATC AAA CAA AAT TGT GAG CTC CTT GAC CAG CTC TCA AGA AAC CTA GCA AAA GTC CCC AGC AAA TGT TGT AAA CAT (1340)  
  
411           420           430           440           450  
tyr thr lys val pro gln val ser thr pro thr leu val glu ser arg asn leu gly lys val ala lys cys cys lys his  
      TAC ACC AAC AAA GTA CCC CAA GTC TCA ACT CCA ACT CTT GTA GAG GTC GTC TCA AGA AAC CTC TTA TGT GTC TGT GAA AAC CCA GTC AGT (1430)  
  
461           468           450           460 461           470  
pro glu ala lys arg met pro cys ala glu asp tyr leu ser val val gln leu cys val leu his glu lys thr dro val ser  
      CCT GAA CCA AAA AGA ATG CCC TGT GCA GAA GAC TAT CTA TCC GTC GTC AAC CAG TCA TAC GAC AAA ACC CCA GTC AGT (1520)  
  
471           476 477           480           490           500  
asp arg val thr lys cys cys thr glu ser leu val asn arg arg pro cys phe ser ala leu glu val asp glu thr tyr val pro lys  
      GAC AGA GTC ACC AAA TGC TGC ACA GAA TCC TGC GTC AAC AGC CGA CCA TGC TTT TCA GCT CTC GAA GTC GAT GAA ACA TAC GTC CCC AAA (1610)  
  
501           510           514           520           530  
glu phe asn ala glu thr phe thr phe his ala asp lle cys thr leu ser glu lys glu arg aln lle lys gln thr ala leu val  
      GAG TTT AAT GCT GAA ACA TTC ACC GTC CAT GCA GAT ATA TGC ACA CTT TGT GAG AAG GAC AGA CAA ATC AAG AAA ACT GCA CTT GTT (1700)

5

10

15

20

25

30

35

0079739

4083

-23-

9. Nucleotide sequence coding for the pre pro human serum albumin, said nucleotide sequence is as follows:

		5		
		10		-10
		15		p r c
		20		Met lys trp val thi phe Ile ser leu leu phe leu ser
		25		ATG AAG TGG GTA ACC TTT ATT TCC CTT CTC TTT AGC (30)
		30		ala tyr ser arg gly val phe arg asp ala his lys ser glu val ala his arg phe lys asp leu ala glu asn phe lys
		35		TCG GCT TAT TCC AGG CGT GTC TTT CGT CGA GAT GAC AGT GAC TGT GCT CAT CGC AAA CAC ACT GCA (170)
		40		ala leu val lle ala phe ala gln tyr leu gln gln cys pro phe glu asp His val lys leu val asn glu val thr glu phe ala
		45		CCC TTC GTC ATT GCC TTT GCT CAG TAT CTT CAG CAG TGT CCA TTT GAA GAT CAT GTA AAA TTA GTG AAT GAA CAA ACT CTT GCA (260)
		50		
		55		lys thr cys val ala asp glu ser ala glu asn cys asp lys ser leu His thr leu phe gly asp lys leu val asn glu val thr leu
		60		AAA ACA TGT GTC CCT GAT GAG TCA GCT GAA ATT TGT GAC AAA TCA CTT CAT ACC CTT TTT CGA GAC ACA ATT TCC ACA (350)
		65		
		70		arg glu thr tyr gly glu met ala asp cys cys ala lys gln glu pro gly arg asn glu cys ala lys asp asn pro
		75		CGT GAA ACC TAT GCT GAA ATG GCT GAC TCC TGT CGA AAA CAA CCT GCA (440)
		80		
		85		ala lys thr cys val met cys thr ala phe his asp asn glu glu thr phe leu cys thr val ala thr leu
		90		100 101
		95		arg glu thr tyr gly glu met ala asp cys cys ala lys gln glu pro gly arg asn glu cys ala lys asp asn pro
		100		CGT GAA ACC TAT GCT GAA ATG GCT GAC TCC TGT CGA AAA CAA CCT GCA (460)
		105		
		110		ala lys leu val pro glu val asp val met cys thr ala phe his asp asn glu glu thr phe leu lys lys asp asn pro
		115		CGT GAA ACC TAT GCT GAA ATG GCT GAC TCC TGT CGA AAA CAA CCT GCA (480)
		120		
		125		ala lys leu val pro glu val asp val met cys thr ala phe his asp asn glu glu thr phe leu lys lys asp asn pro
		130		CGT GAA ACC TAT GCT GAA ATG GCT GAC TCC TGT CGA AAA CAA CCT GCA (500)
		135		
		140		ala lys leu val pro tyro tyr ala pro glu leu phe ala lys ala lys arg tyr lys ala ala phe thr glu cys cys ala
		145		160 169 170
		150		ala lys ala arg his pro tyro tyr ala pro glu leu phe ala lys ala lys arg tyr lys ala ala phe thr glu cys cys ala
		155		CGT GAA ACC TCA GCA ATT GCA GCA (520)
		160		
		165		ala lys ala arg his pro tyro tyr ala pro glu leu phe ala lys ala lys arg tyr lys ala ala phe thr glu cys cys ala
		170		CGT GAA ACC TCA GCA ATT GCA GCA (540)
		175		
		180		ala lys ala ala cys leu leu pro lys leu asp glu alu gly lys ala ser ser als lys ala lys ala lys ala lys cys
		185		CCT GCT GAT AAA GCT GCG TGC CTC CCT GAT GAA CTT CGG GAT GAA GCG AAC GCT TCG TCT GCC AAA CAC ACA CTC AAG TGT (710)
		190		
		195		
		200		
		205		ala ser leu gln lys ala phe gly glu arg ala phe lys als val als arg leu ser gln arg phe pro lys als glu phe als glu
		210		CCC ACT CTC CAA AAA TTT GCA GAA AGA GCT TCC AAA CCT CCC CTC ACC CAC GCA TTT GCA GAA (330)
		215		
		220		
		225		
		230		

0079739

4083

-24-

0079739

4083

-25-

5

10

15

20

25

30

35

531                         540                         550                         558 559 560  
glu leu val lys pro lys ala thr lys glu gln leu lys ala val met asp asp phe ala phe val glu lys cys cys lys  
CAG CTC GTC AAA CAC AAC CCC AAG GCA ACA AAA GAG CAA CTC AAA GCT GTT ATG CAT GAT TTC GCT GCT TTT GTA GAG AAG TGC TCC AAG (1790)

561                         567                         570                         580  
ala asp asp lys glu thr cys phe ala glu glu gln lys leu val ala ser gln ala ala leu gln leu ter ter  
GCT GAC GAT AAC GAG ACC TCC TTT GCC GAG GGT AAA AAA CTT GTC AGT GCA AGT CAA GCT GCC TTA GGC TTA TAA CATCACATTAAAG (1883)

ter                         ter                         ter                         ter

CATCTACGCTTACCATGACAATAAGACAATCAAACCTTATTCATCTGTTTCTTTCTGTTGCTTAAGCCMACACCCCTGCTCTAAAAAACATAAAAAATCTTTAA (2002)

TCACTTTGGCTCTTTCTCTGCTCTCAATTAAATAAAATGAAATCTAA..... 20 .....AA (2078)

10. A nucleotide sequence according to any of claims 6 to 9, in essentially pure form.
11. A DNA transfer vector comprising a nucleotide sequence as defined in claim 5.
- 5 12. A DNA transfer vector according to claim 11, transferred to and replicated in a micro-organism.
13. A DNA transfer vector according to claim 12, which is a plasmid.
14. A DNA transfer vector according to claim 13, 10 wherein the plasmid is pBR322 or YEp6.
15. A process for preparing human serum albumin, which comprises culturing a micro-organism according to claim 5.
16. A DNA transfer vector according to any of 15 claims 12 to 14, or a process according to claim 15, wherein the micro-organism is a bacterium or yeast.
17. A vector or process according to claim 16, wherein the bacterium or yeast is E. coli or Saccharomyces cerevisiae.

0079739

Restriction Endonuclease Map of Human Serum Albumin cDNA Clones

